

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

IN RE:

OPENAI, INC.,
COPYRIGHT INFRINGEMENT LITIGATION

25-md-3143 (SHS) (OTW)
Hon. Sidney H. Stein
Hon. Ona T. Wang

This Document Relates To:

THE NEW YORK TIMES COMPANY v.
MICROSOFT CORPORATION, et al., No. 23-cv-
11195

**OPENAI DEFENDANTS' SUPPLEMENTAL OPPOSITION TO NEWS PLAINTIFFS'
MOTION REGARDING OUTPUT LOGS**

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. PROCEDURAL HISTORY.....	3
A. The Court’s May 13 Order	3
B. OpenAI’s Reasonable, Privacy-Conscious Proposal	4
III. ARGUMENT	5
A. Preservation of <i>All</i> User Conversation Data Is Unprecedented, Unwarranted, and Contrary to Law.....	5
B. Plaintiffs Have Long Been Aware that OpenAI Has Been Complying with its Long-Standing Privacy Policies	12
C. News Plaintiffs Do Not—And Cannot—Justify The Sweeping Preservation Order They Seek.....	12
D. Plaintiffs’ Overreaching Search And Deposition Requests Should Be Rejected.....	13
E. The Court Should Reconsider and Vacate its Preservation Order	15
IV. CONCLUSION	15

TABLE OF AUTHORITIES

	Page(s)
CASES	
<i>Al Otro Lado, Inc. v. Nielsen</i> , 328 F.R.D. 408 (S.D. Cal. 2018)	6
<i>Boudreau v. Petit</i> , No. 17-CV-301-MSM, 2025 WL 553975 (D.R.I. Feb. 19, 2025).....	6
<i>Columbia Pictures Indus. v. Bunnell</i> , No. 06-cv-1093, 2007 WL 2080419 (C.D. Cal. May 29, 2007).....	12, 13
<i>Concord Boat Corp. v. Brunswick Corp.</i> , No. LR-C-95-781, 1997 WL 33352759 (E.D. Ark. Aug. 29, 1997).....	10, 11
<i>Concord Music Grp., Inc. v. Anthropic PBC</i> , No. 24-cv-03811-EKL (N.D. Cal. Mar. 25, 2025)	14
<i>Gill v. Gulfstream Park Racing Ass'n., Inc.</i> , 399 F.3d 391 (1st Cir. 2005).....	8
<i>Gonzales v. Google, Inc.</i> , 234 F.R.D. 674 (N.D. Cal. 2006).....	8, 9
<i>Griffith v. TikTok, Inc.</i> , No. 5:23-cv-00964-SB-E, 2024 WL 4875339 (C.D. Cal. Oct. 10, 2024)	11
<i>Griffith v. TikTok, Inc.</i> , No. 5:23-cv-00964-SB-E, 2024 WL 5279224 (C.D. Cal. Dec. 24, 2024).....	11
<i>In re Application of Chevron Corp.</i> , 736 F. Supp. 2d 773 (S.D.N.Y. 2010).....	10
<i>In re Sealed Case (Med. Recs.)</i> , 381 F.3d 1205 (D.C. Cir. 2004).....	10
<i>Ottoson v. SMBC Leasing & Fin., Inc.</i> , 268 F. Supp. 3d 570 (S.D.N.Y. 2017).....	13
<i>Riley v. California</i> , 573 U.S. 373 (2014).....	7, 8
<i>Simmons v. Danhauer & Assocs., LLC</i> , No. 08-cv-3819, 2009 WL 10677391 (D.S.C. Sept. 15, 2009)	13

Treppel v. Biovail Corp.,
233 F.R.D. 363 (S.D.N.Y. 2006) 6

Viacom Int'l Inc. v. Youtube Inc.,
253 F.R.D. 256 (S.D.N.Y. 2008) 12

Zubulake v. UBS Warburg LLC,
220 F.R.D. 212 (S.D.N.Y. 2003) 5, 11

RULES

Fed. R. Civ. P. 37 6

Fed. R. Civ. P. 37(e) 5, 11, 13

Rule 26(b)(1) 6

Rule 30(b)(6) 14

OTHER AUTHORITIES

OpenAI, Terms of Use, <https://openai.com/policies/terms-of-use> 7

OpenAI, Privacy Policy, <https://openai.com/policies/row-privacy-policy/> 1, 7

I. INTRODUCTION

“Many Americans now rely on digital tools to work remotely and stay connected. They shouldn’t have to sacrifice their privacy to use them.”¹ The New York Times Editorial Board wrote these words just a few years ago, and they reflect a guiding principle for how OpenAI operates today. OpenAI is strongly committed to the privacy of its users and to keeping its users’ information secure. For this reason, OpenAI recognizes that users own their conversations (both the user prompts and outputs) and empowers them to delete these conversations so that they can protect their sensitive or personal information. Yet News Plaintiffs—including The New York Times—are now seeking to force OpenAI to disregard its users’ decisions about which of their ChatGPT conversations can be deleted, all so News Plaintiffs might gain tactical leverage in a commercial dispute. OpenAI opposes this unprecedented attempt to undermine the privacy rights of ChatGPT users and coerce OpenAI into violating its commitments to them.² Indeed, as The New York Times has acknowledged, “granting users some semblance of control over their own data is more urgent than ever.”³ OpenAI agrees. The New York Times and other News Plaintiffs’ effort to deprive OpenAI users of their ability to delete their data is hypocritical, and it is wrong.

While nothing could justify such a blatant affront to user privacy, News Plaintiffs’ stated rationale is particularly weak. In effect, News Plaintiffs ask the Court to override the decisions of OpenAI’s users in order to facilitate a dragnet for evidence there is no reason to believe exists, much less in any material amount. News Plaintiffs themselves (over)estimate that only 0.006% of the data might even be relevant. Nightingale Dawson Decl. Ex. A. News Plaintiffs’ overreaching

¹ The New York Times Editorial Board, *Privacy Cannot Be a Casualty of the Coronavirus*, N.Y. Times (Apr. 7, 2020), <https://www.nytimes.com/2020/04/07/opinion/digital-privacy-coronavirus.html>.

² See OpenAI, *Privacy Policy*, <https://openai.com/policies/row-privacy-policy/>.

³ The New York Times Editorial Board, *America, Your Privacy Settings Are All Wrong*, N.Y. Times (Mar. 6, 2021), <https://www.nytimes.com/2021/03/06/opinion/data-tech-privacy-opt-in.html>.

demand is akin to requiring Google to block all Gmail users from deleting any emails because of the mere speculation (or remote possibility) that a search might uncover instances of users sending messages that copy some or all of a copyrighted article. Such an outcome is unprecedented, and it would set a dangerous precedent. The reasoning offered by News Plaintiffs for why this extreme remedy is needed rests entirely on the false premise that (i) ChatGPT is used to bypass paywalls and (ii) that these users are more likely to delete their conversations than others. ChatGPT is not designed to bypass paywalls of news sites; in fact, the opposite is true. Monaco Decl. ¶ 9. And although News Plaintiffs have alleged that they could generate snippets of their material in ChatGPT by manipulating prompts and violating OpenAI's terms of service—including threatening “thermonuclear war” and the “life of a kitten” unless ChatGPT complied⁴—there is nothing to suggest that anyone else has misused OpenAI’s products in the same manner.⁵ Regardless, it is entirely speculative to suggest the conversations users have chosen to delete are different from those that have been retained.⁶ OpenAI has proposed a path forward that would provide for the preservation of user conversations where there is evidence of relevance, as well as a mechanism to show that there is no relevant difference between deleted conversations and other conversations, with the parties agreeing to seek a vacatur of the preservation order, but News Plaintiffs have rejected it.

Every day, millions of people use ChatGPT to assist with everyday tasks and to improve their lives, from planning vacations to developing business ideas to serving as a sounding board to

⁴ See, e.g., Exhibit 10 to CIR Am. Compl. (*Center for Investigative Reporting*, No. 1:24-cv-04872, ECF 88-13) at 1; Exhibit J to NYT Am. Compl. (NYT ECF 202-1) (manipulating old version of OpenAI model to generate output referring to non-existent Times article); *New York Times*, No. 1:23-cv-11195, Sep. 12, 2024 Hearing Tr. at 58 (Times’ counsel admitting that the “purpose of Exhibit J is not to demonstrate how your average user would use GPT”).

⁵ Indeed, while seeking overreaching and speculative discovery on OpenAI’s users on the one hand, News Plaintiffs on the other hand have refused any discovery into the creation of those known, specific coerced examples, presumably because they know the examples are not the product of ordinary use of OpenAI’s services but of News Plaintiffs’ intentional manipulation to generate evidence to support their claims.

⁶ OpenAI conferred with Plaintiffs, per the Court’s order (NYT ECF 559), but was unable to resolve the dispute.

express their most private thoughts, feelings, and concerns. OpenAI is strongly committed to user privacy and to permitting users to decide which of their conversations to delete and which to keep. The issue before the Court in this dispute is not about the actions of OpenAI—which has not itself decided to “destroy” conversations—but whether OpenAI’s users should have a say in what happens to their data. The New York Times’ Editorial Board was right that “consumers need to have control over their personal information.”⁷ The New York Times’ lawyers are wrong to try to take away that control, and this Court should not let them do so. The Court’s Order should be vacated and News Plaintiffs’ motion denied.

II. PROCEDURAL HISTORY

A. The Court’s May 13 Order

On January 13, 2025, The New York Times and Daily News Plaintiffs moved the Court for an order requiring wholesale preservation of users’ written prompts and outputs (hereafter “Conversation Data”) regardless of whether a user has chosen not to retain that data. *See NYT ECF 379.* OpenAI opposed, explaining that the request for the preservation of irrelevant information is unduly burdensome and disproportionate to the needs of this case. *NYT ECF 423.* The Court denied the request for “wholesale preservation,” and directed the parties to confer further on the “the privacy and technological considerations implicated in Plaintiffs’ request.” *NYT ECF 441.* Since then, OpenAI has conferred repeatedly with News Plaintiffs regarding both this dispute and Conversation Data more generally. *See* Nightingale Dawson Decl. ¶ 4. But The New York Times and Daily News Plaintiffs, now joined by Center for Investigative Reporting (“CIR”)⁸ then filed—without warning—a supplemental letter on May 12, *NYT ECF 546* (which

⁷ The New York Times Editorial Board, *Protecting the Privacy of Internet Users*, N.Y. Times (Mar. 11, 2016), <https://www.nytimes.com/2016/03/12/opinion/protecting-the-privacy-of-internet-users.html>.

⁸ This brief refers to The New York Times, the Daily News Plaintiffs, and CIR collectively as “News Plaintiffs.”

was replete with factual inaccuracies, *see NYT ECF 558*). Subsequently, this Court issued a preservation order regarding Conversation Data. *NYT ECF 551*. Soon after the order was issued, ChatGPT users and other third parties raised concerns about the impact of the order on their privacy. *See Nightingale Dawson Exs. C–L*. OpenAI promptly sought reconsideration of the preservation order, *NYT ECF 558*, which the Court denied without prejudice to renewal in this submission. *NYT ECF 559*.

B. OpenAI’s Reasonable, Privacy-Conscious Proposal

Per the Court’s directive, *NYT ECF 559* at 5, as part of the parties’ ongoing meet-and-confer efforts regarding this dispute, OpenAI has proposed a resolution to appropriately scope any additional preservation of Conversation Data to balance the privacy rights of OpenAI’s customers with News Plaintiffs’ ability to obtain relevant evidence (if any).

First, OpenAI proposed that News Plaintiffs provide OpenAI with evidence sufficient to identify users who have employed OpenAI’s products to generate outputs that News Plaintiffs believe are relevant. OpenAI would then implement appropriate going-forward preservation for any associated accounts’ Conversation Data. To date, News Plaintiffs have not identified any such evidence. *See Nightingale Dawson Decl.* ¶ 5.

Second, OpenAI proposed a mechanism to investigate “whether the retained output logs marked for deletion differ in any material way from the retained logs,” *NYT ECF 559* at 5, and recommended that—to avoid needless burden on OpenAI and its users—the parties stipulate to vacatur of the Court’s preservation order given OpenAI’s proposal. *Nightingale Dawson Decl.* ¶ 5. In particular, OpenAI referenced its earlier proposals to provide News Plaintiffs a process for sampling ChatGPT Conversation Data⁹ for analysis by the parties’ respective experts. *Id.* As part

⁹ “ChatGPT Conversation Data” refers to Conversation Data for ChatGPT Free, Plus, and Pro.

of that process, OpenAI offered to make available two anonymized, random samples of ChatGPT Conversation Data: the first representing conversations that OpenAI ordinarily retains, and the second representing conversations that OpenAI ordinarily retains *plus* conversations that would otherwise not be retained under its policies. *Id.* Those two samples would allow News Plaintiffs to assess whether (as they speculate) the “prevalence of harmful outputs is higher in the destroyed data.” *NYT* ECF 560 at 2. OpenAI agreed to meet-and-confer with Plaintiffs to address whether, given the results of such analysis, any further appropriately scoped preservation is necessary. Nightingale Dawson Decl. ¶ 5. This approach would strike an appropriate and reasonable balance between the competing interests and expectations of OpenAI’s users and News Plaintiffs’ litigation aims. Regrettably, the News Plaintiffs rejected OpenAI’s proposal. *Id.*

III. ARGUMENT

What News Plaintiffs seek from this Court is unprecedented: a preservation order requiring a company to preserve *all* private user data, without evidence of relevance, and in direct contravention of its users’ decisions regarding their private information. News Plaintiffs have not identified a single example of a court issuing such a sweeping order, and OpenAI is aware of none. The privacy interests of non-party users of OpenAI’s products, along with the significant harm wholesale preservation imposes on OpenAI, significantly outweigh the (entirely speculative) benefit News Plaintiffs claim to seek. Wholesale preservation should be denied.

A. Preservation of *All* User Conversation Data Is Unprecedented, Unwarranted, and Contrary to Law

A duty to preserve does not require preserving “every e-mail or electronic document.” *Zubulake v. UBS Warburg LLC*, 220 F.R.D. 212, 217 (S.D.N.Y. 2003) (citation omitted). Instead, preservation requirements must be “reasonable.” See Fed. R. Civ. P. 37(e). And a key “factor in

evaluating the reasonableness of preservation efforts is proportionality.”¹⁰ This approach makes sense as “the scope of a party’s duty to preserve is the same as the scope of discovery articulated in Rule 26(b)(1).” *Al Otro Lado*, 328 F.R.D. 408 at 416. At a minimum, a duty to preserve does not arise absent some showing that the “documents in jeopardy are in fact relevant.” *Treppel v. Biovail Corp.*, 233 F.R.D. 363, 370–71 (S.D.N.Y. 2006).

Here, News Plaintiffs’ demand for indiscriminate, wholesale preservation of all private Conversation Data—*against users’ express decisions to delete such data*—is the opposite of reasonable and is contrary to law. News Plaintiffs have made no attempt to account for the harms that flow to OpenAI *and* its users from wholesale preservation, let alone to balance those harms against the (lack of) probative value of the Conversation Data they want preserved. There is *no* evidentiary basis to conclude that the additional Conversation Data News Plaintiffs want preserved has any probative value at all beyond the many billions of already retained examples that News Plaintiffs still have not analyzed. Simply put, wholesale preservation of Conversation Data is unjustified because (1) it undermines users’ privacy rights and erodes users’ trust; (2) it requires OpenAI to expend significant technical and financial resources, including to create and maintain new technological infrastructure; and (3) it is based on pure speculation by News Plaintiffs.

1. *OpenAI’s Users’ Privacy Interests Deserve Respect and Consideration*

ChatGPT users’ right to privacy—and control of their data—is a fundamental principle; it underpins the trust and reliance millions of individuals place in OpenAI daily. Kechida Decl. ¶ 2. That is why OpenAI’s users—not OpenAI—own their Conversation Data, both the prompts and

¹⁰ Fed. R. Civ. P. 37 advisory committee’s notes to 2015 amendment; *see also Boudreau v. Petit*, No. 17-CV-301-MSM, 2025 WL 553975, at *2 (D.R.I. Feb. 19, 2025) (noting that preservation obligation should balance the burden of preservation against the contents of documents sought to be preserved); *Al Otro Lado, Inc. v. Nielsen*, 328 F.R.D. 408, 416 (S.D. Cal. 2018) (“Preservation is evaluated on a standard of reasonableness, which is in turn informed by proportionality.”).

outputs. *See, e.g.*, OpenAI, Terms of Use, <https://openai.com/policies/terms-of-use>. And it is why OpenAI has established clear policies that respect and protect this right, offering users the ability to keep or delete their private Conversation Data.¹¹ *See* Kechida Decl. ¶ 3. Those policies foster an environment where users can engage freely and openly, and make decisions about their personal data. Indeed, OpenAI has done exactly what plaintiff The Chicago Tribune said “tech companies” *should* do: have “the public’s best interests in mind when it comes to protecting personal data” and “handle data responsibly.”¹² Doing so, as The Chicago Tribune recognized then (but apparently no longer thinks now), is critical to protect ChatGPT users’ privacy interests.

Like the “cache of sensitive personal information” on a cell phone or an “Internet search and browsing history,” ChatGPT Conversation Data range “from the mundane to the intimate.” *Riley v. California*, 573 U.S. 373, 395 (2014). People use ChatGPT to create household budgets, providing details about everything from their income to their expenses to their money-saving goals.¹³ They use ChatGPT to write wedding vows, sharing intimate details of their relationships.¹⁴ They use ChatGPT to generate gift ideas, sharing tidbits about their loved ones.¹⁵ And they use ChatGPT to create business plans, sharing highly confidential and proprietary business information.¹⁶ People do all of that and more because they know OpenAI is committed to protecting their privacy and they have control of their Conversation Data.

¹¹ OpenAI, *Privacy Policy*, <https://openai.com/policies/row-privacy-policy>.

¹² The Chicago Tribune Editorial Board, *A Thumb’s Up For Illinois Digital Privacy Law*, The Chicago Tribune, <https://www.chicagotribune.com/2019/02/01/a-thumbs-up-for-illinois-digital-privacy-law/> (Feb. 2, 2019).

¹³ *See* Karen Bennett, *How to make a budget using ChatGPT*, Bankrate, <https://www.bankrate.com/personal-finance/how-to-make-budget-with-chatgpt/#list-out-expenses> (Feb. 10, 2025).

¹⁴ Samantha Kelly, *Who says romance is dead? Couples are using ChatGPT to write their wedding vows*, CNN Business, <https://www.cnn.com/2023/04/12/tech/chatgpt-wedding-vows/index.html> (Apr. 12, 2023).

¹⁵ Amanda Caswell, *7 ChatGPT prompts I’m using to improve my gift giving — here’s how to try them*, tom’s guide, <https://www.tomsguide.com/ai/chatgpt/7-chatgpt-prompts-im-using-to-improve-my-gift-giving-heres-how-to-try-them> (Dec. 3, 2024).

¹⁶ Aytokin Tank, *The Entrepreneur’s Guide To Writing A Winning Business Plan With ChatGPT*, Forbes, <https://www.forbes.com/sites/aytekintank/2024/09/17/9-chatgpt-prompts-to-write-a-business-plan/> (Sept. 17, 2024).

When determining whether something is discoverable, which informs whether it can be subject to a preservation order, “considerations of the public interest, the need for confidentiality, and privacy interests are relevant factors to be balanced.” *Gill v. Gulfstream Park Racing Ass’n., Inc.*, 399 F.3d 391, 402 (1st Cir. 2005); *see also Gonzales v. Google, Inc.*, 234 F.R.D. 674, 687 (N.D. Cal. 2006). Those privacy interests attach to a broad range of user data, much like those at issue here. *See Gonzales*, 234 F.R.D. at 687 (rejecting Government’s position that “the mere text of [search] queries would not yield identifiable information”); *Riley*, 573 U.S. at 395. In a similar case, the court raised these concerns *sua sponte*, expressing deep concern for the “serious privacy concerns” raised by the use of users’ internet search queries in litigation. *Gonzales*, 234 F.R.D. at 687 (noting the difference between “a user’s search query reading ‘[user name] stanford glee club’” and “a user’s search for ‘[user name] third trimester abortion san jose’”).

The privacy interests of OpenAI’s users—non-parties to this litigation—heavily favor rejecting News Plaintiffs’ demand for wholesale preservation. User privacy interests are at their apex regarding Conversation Data users choose to delete, whether by deleting a single ChatGPT message or deleting an account—with the expectation that their private information will be treated pursuant to OpenAI’s commitments to them, reflecting industry-wide standards and norms. This choice is a fundamental expression of their privacy rights, allowing individuals to control their own data and ensuring that their most intimate inquiries remain confidential. In a digital age where data privacy is increasingly under threat, safeguarding these user-directed deletions is crucial.

This Court need not take just OpenAI’s word for it. Earlier this week, ChatGPT user Spark Innovations Corp. filed a motion to intervene “for the limited purpose of objecting to the retention and preservation” of its “confidential, trade-secret, and competitively sensitive business information,” premised on News Plaintiffs “overreaching into the private affairs of third parties

with no stake in this litigation.” *NYT ECF* 566. Likewise, press coverage has referred to the danger of letting users’ privacy “fall by the wayside without a second thought rather than impose any kind of duty on the plaintiffs to show the retention of this data would advance the litigation in a meaningful way.”¹⁷ As one legal commentator noted, this could “shake the trust of consumers who expected their chatbot records to be deleted.” *Id.* Numerous other users and lawyers alike have raised these same concerns in various online forums. *Nightingale Dawson Decl.* Exs. C–L. As the court in *Gonzales* recognized, safeguarding user privacy has an “appreciable impact” on the relationship between a company and its users. *Gonzales*, 234 F.R.D. at 684.

2. *News Plaintiffs’ Wholesale Preservation Demands Will Impose Significant Burdens on OpenAI*

As a technical matter, although OpenAI has taken steps immediately available to it to comply with the order—such as provisionally retaining in a segregated legal hold database certain ChatGPT Conversation Data—its ongoing investigation confirms that expanding that wholesale ongoing preservation to include all Conversation Data for all users, and for all products, will require OpenAI to undertake a massive project to overhaul and rebuild elements of its core data infrastructure. *Monaco Decl.* ¶ 8. OpenAI estimates that such an undertaking will require, at least, months of engineering time along with additional costs in the [REDACTED] of dollars. *Id.* These are resources that could otherwise be used for mission-critical functions like ongoing research, deployment of safe and beneficial AI, model safety and alignment, and product development.

Moreover, wholesale preservation complicates OpenAI’s ongoing efforts to comply with data privacy laws and regulations around the world, many of which specifically bear on retention of user data and users’ right to delete it. *See Kechida Decl.* ¶¶ 2–4. Worse still, wholesale

¹⁷ Kathryn Lundstrom, *A Federal Judge Ordered OpenAI to Stop Deleting Data. Here’s How That Could Impact Users’ Privacy*, Adweek (May 21, 2025) <https://www.adweek.com/media/a-federal-judge-ordered-openai-to-stop-deleting-data-heres-how-that-could-impact-users-privacy>.

preservation threatens OpenAI’s ongoing ability to meet its commitments to users, many of whom prioritize privacy when choosing between technology platforms. Many of OpenAI’s users have already reacted to news about the Court’s May 13 Order with significant concern. *See* Nightingale Dawson Decl. Exs. C–L. News Plaintiffs’ demand threatens the company’s commitments to its users, extending beyond technical burden.

3. *News Plaintiffs Have No Basis to Speculate That Users Delete Private Conversation Data That Is Relevant*

Courts “should not” compel parties to “preserve potentially relevant evidence . . . in the absence of some substantial basis for believing that it would serve a useful purpose.” *In re Application of Chevron Corp.*, 736 F. Supp. 2d 773, 787 (S.D.N.Y. 2010). Instead, when considering whether data should be subject to a wholesale preservation order, courts should weigh—as is necessary to determine whether production should be required, *see In re Sealed Case (Med. Recs.)*, 381 F.3d 1205, 1217–18 (D.C. Cir. 2004)—the “privacy interest in a record against the record’s value.” *Id.*

In this case, there is no basis—let alone a “substantial” one—to believe that a wholesale preservation order, such as the one-at-issue, would “serve a useful purpose.” *Id.* News Plaintiffs themselves implicitly concede that 99.994% of the Conversation Data that they demand OpenAI preserve will not “implicate [their] claims” at all, and have provided no basis for the remaining, speculative 0.006%. Nightingale Dawson Decl. Ex. A. Although News Plaintiffs assert that deleted Conversation Data could hold heightened relevance, that assertion should be taken for what it is: pure speculation. *Concord Boat Corp. v. Brunswick Corp.*, No. LR-C-95-781, 1997 WL 33352759 (E.D. Ark. Aug. 29, 1997) (questioning “inferential leap” that “deleted e-mails must have been relevant”). Their argument rests entirely on News Plaintiffs’ own contrived manipulations of OpenAI’s products, resulting in the creation of, for example, Exhibit J to The

New York Times's complaint. *See supra* n. 1.¹⁸ News Plaintiffs have not identified a single instance of a non-party ChatGPT user employing it to access News Plaintiffs' articles behind paywalls or to obtain copies of such articles and then to somehow reproduce those articles in Conversation Data, let alone evidence of someone doing so and then deleting their Conversation Data; this failure of proof is fatal to News Plaintiffs' request. *Cf. Griffith v. TikTok, Inc.*, No. 5:23-cv-00964-SB-E, 2024 WL 5279224, at *8 (C.D. Cal. Dec. 24, 2024) ("Plaintiffs have not shown any basis for the Court to infer that the deleted data contained evidence necessary to support their claims that is materially different from the data in the sample.").

* * *

The Federal Rules contemplate that a party take "reasonable steps" to preserve data that are proportionate to the needs of the case. *See Fed. R. Civ. P. 37(e); id.* advisory committee's notes to 2015 amendment. News Plaintiffs' demand for wholesale preservation requires unreasonable—and unduly burdensome—steps that are not proportionate to the needs of the case. News Plaintiffs' approach—that one must preserve all the haystacks in the world based on the most remote and unsupported possibility that one might contain a needle—would "cripple" companies by requiring them to preserve everything. *See Zubulake*, 220 F.R.D. at 217. That is why News Plaintiffs' approach has been rejected by the Federal Rules and federal courts alike. *See, e.g., id.; Griffith v. TikTok, Inc.*, No. 5:23-cv-00964-SB-E, 2024 WL 4875339, at *3 (C.D. Cal. Oct. 10, 2024) (declining to impose sanctions when company did not change two-week retention period given the "millions of dollars" in cost, "privacy and security concerns about preserving the data," and fact that defendant offered sampled data).

¹⁸ Daily News Plaintiffs have also alleged the existence of so-called "pink-slime" news websites (*Daily News*, ECF 1 ¶ 146), but the only factual support they cite is an article discussing a contrived exercise, not any real-world examples (*see id.* n. 55) or evidence that OpenAI's models (as opposed to other language models) generated the "pink-slime" websites of which they complain.

B. Plaintiffs Have Long Been Aware that OpenAI Has Been Complying with its Long-Standing Privacy Policies

In February 2024, OpenAI informed counsel for The New York Times, who is also counsel for the Daily News Plaintiffs, that it would be continuing to comply with its longstanding—and publicly available—privacy policies and that it would be “inappropriate, overly burdensome, not proportional, and not feasible to place all ChatGPT inputs and output on indefinite retention.” *See* *NYT* ECF 379-3. OpenAI’s privacy policies and retention practices were no secret. Indeed, CIR acknowledged these exact privacy policies and retention practices in its June 2024 Complaint. *CIR* ECF 1 ¶ 77 (citing OpenAI’s publicly available retention policy). And when The Times became an OpenAI customer in June 2023, its agreement with OpenAI prevented OpenAI from logging, saving, or retaining certain Times Conversation Data. *See* Nightingale Dawson Decl. Ex. B.

Belying any claim of purported relevance or importance, much less any urgency, The Times never even responded to OpenAI’s February 2024 letter. Nightingale Dawson Decl. ¶ 6. The first time The Times or the Daily News raised an objection to OpenAI’s data practices was 10 months later in December 2024. *NYT* ECF 379-14. And CIR—which discussed in their June 2024 complaint OpenAI’s retention policies, *CIR* ECF 1 ¶ 77—has never raised an objection until now.

C. News Plaintiffs Do Not—And Cannot—Justify The Sweeping Preservation Order They Seek

News Plaintiffs rely on three cases for the proposition that this Court should disregard OpenAI’s privacy practices. Each case, however, either supports OpenAI or is distinguishable (or both). In *Viacom Int’l Inc. v. YouTube Inc.*, for example, the court refused the plaintiffs’ request for production of private YouTube videos—that is, the private user content directly analogous to the Conversation Data at issue here. *See* 253 F.R.D. 256, 264-65 (S.D.N.Y. 2008). In *Columbia Pictures Indus. v. Bunnell*, No. 06-cv-1093, 2007 WL 2080419 (C.D. Cal. May 29, 2007), the court explicitly found that retention would “not be unduly burden[some],” *id.* at *8, unlike here, where

compliance would impose significant monetary and non-monetary burdens on OpenAI, *see Monaco Decl.* ¶¶ 4–5. Finally, *Simmons v. Danhauer & Assocs., LLC*, is inapposite, as it addressed a company’s client list, not any client’s personal conversations, such as those at issue here. No. 08-cv-3819, 2009 WL 10677391, at *2 (D.S.C. Sept. 15, 2009). In short, none of these cases justify the wholesale retention—without *any* evidentiary support for the alleged relevance of the data at issue—that News Plaintiffs urge.¹⁹

D. Plaintiffs’ Overreaching Search And Deposition Requests Should Be Rejected

Search. News Plaintiffs are well aware that their proposed search of *all* OpenAI user Conversation Data is technically infeasible. OpenAI has no “preexisting technological functionality to . . . somehow [identify and] preserve only user conversations that reference Plaintiffs’ content.” *NYT* ECF 423 at 2; *see also* Monaco Decl. ¶ 5.²⁰ As an initial matter, OpenAI would need to spend more than [REDACTED] in costs alone just to decompress the current volume of Conversation Data into a searchable format, and over [REDACTED] per month to store it – these numbers would of course increase as the volume of the data increases. Monaco Decl. ¶ 4. Even then, as has been previously explained to News Plaintiffs no feasible method currently exists to search the uncompressed data for all potentially responsive Conversation Data. *See* E. Nightingale Dawson 3/31/2025 Letter. Indeed, in *Concord Music Grp., Inc. v. Anthropic PBC*, on which News

¹⁹ Plaintiffs allege that “spoliation sanctions may be warranted.” Plts. Br. at 1. This is frivolous. To justify sanctions, Plaintiffs would bear the burden of proving not only that (1) deleted Conversation Data “should have been preserved,” (2) the deletion resulted from a failure “to take reasonable steps,” (3) Plaintiffs suffered “prejudice” as a result, and (4) OpenAI “acted with the intent to deprive another party of the information’s use in the litigation.” Fed. R. Civ. P. 37(e); *Ottoson v. SMBC Leasing & Fin., Inc.*, 268 F. Supp. 3d 570, 580 (S.D.N.Y. 2017). News Plaintiffs cannot make any of those showings—there is no evidence that the deleted Conversation Data should have been preserved, OpenAI did take reasonable steps, Plaintiffs have suffered no prejudice as a result of the deletions, and, perhaps most importantly, OpenAI acted with the intent of *ensuring* Plaintiffs knew exactly what OpenAI was doing in light of its privacy policies—the exact opposite of what is required for spoliation sanctions. There is thus no basis upon which Plaintiffs could seek spoliation sanctions, which is presumably why they have not done so.

²⁰ As detailed in the Monaco Declaration, the fact that OpenAI can search the last 30 days’ worth of ChatGPT conversations does not mean that OpenAI can also search “the full extant output log data” for ChatGPT, which is stored in a different (and non-searchable) location. Monaco Decl. ¶ 5.

Plaintiffs rely (Plts. Br. at n.4), the court rejected a proposal to search six months' worth of Anthropic's user prompt and output data, finding that "searches over the entire dataset present issues of technical infeasibility and burden." No. 24-cv-03811-EKL, ECF 318, slip op. at 2 (N.D. Cal. Mar. 25, 2025). The same is even more true here, because unlike in *Anthropic*, where "nearly a million" user conversation records were at issue, News Plaintiffs' proposed search in this case would involve *over 60 billion* such records. *Id.*, ECF 302, slip op. at 7 (N.D. Cal. Mar. 5, 2025); *NYT* ECF 560-18.²¹

Deposition. News Plaintiffs do not identify any basis for their request for yet another standalone, piecemeal Rule 30(b)(6) deposition regarding discovery, and there is none. Even were such discovery relevant, News Plaintiffs can seek it in a Rule 30(b)(6) deposition in the ordinary course. Rather than justify their deposition demand, News Plaintiffs instead suggest that the parties' ongoing conferral efforts—and meeting with OpenAI's technical expert—were somehow unavailing. What Plaintiffs fail to disclose is the voluminous information OpenAI has already provided, both in the form of written responses to Plaintiffs' many letters as well as during the meeting with OpenAI's technical expert, during which he responded to *over 50* highly detailed, substantive technical questions. Nightingale Dawson Decl. ¶ 4. Consistent with OpenAI's good-faith efforts to engage with Plaintiffs to date, OpenAI has also already provided answers, where available, to the questions on which Plaintiffs seek a deposition. See *NYT* ECF 539-1, Ex. A (explaining that it is infeasible to search for News Plaintiffs' asserted works within the billions of

²¹ News Plaintiffs contend that OpenAI can somehow search all of the Conversation Data for their asserted works. (Pls' Brief at 5.) That is wrong. OpenAI was able to run a limited search for user accounts related to this litigation (based on the specific allegations in News Plaintiffs' complaints) in order to comply with its legal preservation obligations of retaining relevant data, despite News Plaintiffs refusal to provide information to support the process. These were however simple text queries of a few sentences on the last 30 days of ChatGPT Conversation Data. This approach cannot practicably be used to search for the full scope of News Plaintiffs' asserted works in this small slice of ChatGPT Conversation Data—let alone in the full dataset.

user conversations OpenAI retained); *NYT* ECF 491 (explaining that Conversation Data that OpenAI has been deleted at the direction of its users cannot be retrieved); *NYT* ECF 558 (explaining the [REDACTED]
[REDACTED]); *NYT* ECF 539-1, Ex. A (detailing the volume of ChatGPT Conversation Data that users have deleted or designated as temporary); *see also* Monaco Decl. ¶¶ 4-5, 10-14.

E. The Court Should Reconsider and Vacate its Preservation Order

To correct a manifest injustice and unjustified intrusion on user privacy, OpenAI renews its request that the Court reconsider its May 13 Order in light of the arguments and declarations attached hereto that the Court did not have the benefit of considering during its initial order. While the Court should reject News Plaintiffs' demand outright, OpenAI is also ready to proceed with its reasonable, privacy-conscious proposal. *See supra*, Section II(B). As detailed above, there is no basis to order ongoing, wholesale preservation of Conversation Data based on the mere speculation that entirely improbable data might exist—particularly where there are *billions* of retained conversations from which to sample. OpenAI respectfully requests that the Order be vacated.

The Court should likewise reconsider and vacate its order in light of OpenAI's proposed path forward pursuant to which (a) Plaintiffs identify any specific accounts for which there is evidence that retention is appropriate, and (b) Plaintiffs are able to engage in additional sampling of anonymized Conversation Data to address Plaintiffs' and the Court's questions while at the same protecting the privacy rights of OpenAI's users.

IV. CONCLUSION

For the foregoing reasons, OpenAI respectfully requests that the Court deny News Plaintiffs' requests and vacate the May 13, 2025 Order at *NYT* ECF 551.

Dated: May 23, 2025

/s/ Elana Nightingale Dawson

LATHAM & WATKINS LLP

Andrew M. Gass (*pro hac vice*)

andrew.gass@lw.com

505 Montgomery Street, Suite 2000

San Francisco, CA 94111

Telephone: 415.391.0600

Sarang V. Damle

sy.damle@lw.com

Elana Nightingale Dawson (*pro hac vice*)

elana.nightingaledawson@lw.com

555 Eleventh Street, NW, Suite 1000

Washington, D.C. 20004

Telephone: 202.637.2200

Rachel R. Blitzer

rachel.blitzer@lw.com

Herman H. Yue

herman.yue@lw.com

Luke A. Budiardjo

luke.budiardjo@lw.com

1271 Avenue of the Americas

New York, NY 10020

Telephone: 212.906.1200

Allison S. Blanco (*pro hac vice*)

allison.blanco@lw.com

650 Town Center Drive 20th Floor

Costa Mesa, CA 92626

Telephone: 714.540.1235

/s/ Joseph C. Gratz

MORRISON & FOERSTER LLP

Joseph C. Gratz (*pro hac vice*)^{*}

jgratz@mofo.com

Rose S. Lee (*pro hac vice*)

roselee@mofo.com

707 Wilshire Boulevard, Suite 6000

Los Angeles, CA 90017

Telephone: 213.892.5200

Carolyn M. Homer (*pro hac vice*)

cmhomer@mofo.com

2100 L Street, NW, Suite 900

Washington, DC 20037

Telephone: 202.650.4597

Eric K. Nikolaides

enikolaides@mofo.com

Emily C. Wood

ewood@mofo.com

Jocelyn E. Greer

jgreer@mofo.com

250 West 55th Street

New York, NY 10019

Telephone: 212.468.8000

Attorneys for OpenAI Defendants

* All parties whose electronic signatures are included herein have consented to the filing of this document, as contemplated by Rule 8.5(b) of the Court's ECF Rules and Instructions.

/s/ Edward A. Bayley

KEKER, VAN NEST & PETERS
LLP Robert A. Van Nest (*pro hac vice*)
rvannest@keker.com
Paven Malhotra
pmalhotra@keker.com
Michelle S. Ybarra (*pro hac vice*)
mybarra@keker.com
Nicholas S. Goldberg (*pro hac vice*)
ngoldberg@keker.com
Thomas E. Gorman (*pro hac vice*)
tgorman@keker.com
Katie Lynn Joyce (*pro hac vice*)
kjoyce@keker.com
Sarah Salomon (*pro hac vice*)
ssalomon@keker.com
R. James Slaughter (*pro hac vice*)
rslaughter@keker.com
Christopher S. Sun (*pro hac vice*)
csun@keker.com
Andrew F. Dawson (*pro hac vice*)
adawson@keker.com
Andrew S. Bruns (*pro hac vice*)
abrunz@keker.com
Edward A. Bayley (*pro hac vice*)*
ebayley@keker.com
Ryan K. Wong (*pro hac vice*)
rwong@keker.com
Maile N. Yeats-Rowe (*pro hac vice*)
myeats-rowe@keker.com
Spencer McManus (*pro hac vice pending*)
smcmanus@keker.com
Olivia C. Malone (*pro hac vice pending*)
omalone@keker.com
633 Battery St.
San Francisco, CA 94111
Telephone: 415.391.5400

Attorneys for OpenAI Defendants